AD-A285 904

Technical Report 1009

Predicting Table VIII Tank Gunnery Performance From M-COFT Hit Rate

Monte D. Smith
CAE-Link Corporation

Joseph D. Hagman
U.S. Army Research Institute

October 1994

ARMY RESEARCH INSTITUTE

United States Army Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract for the Department of the Army

CAE-Link Corporation

Technical review by

J. Douglas Dressel
Ronald E. Kraemer

Accesion For	
NTIS CRA&I	✓
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
1994, October

3. REPORT TYPE AND DATES COVERED
Final Jun 93 - Mar 94

4. TITLE AND SUBTITLE
Predicting Table VIII Tank Gunnery Performance From M-COFT Hit Rate

5. FUNDING NUMBERS
DAAL03-91-C-0034
63007A
793
2125
C01

6. AUTHOR(S)
Smith, Monte D. (CAE-Link); and Hagman, Joseph D. (ARI)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CAE-Link Corporation
Link Training Services Division
5111 Leesburg Pike, Suite 300
Falls Church, VA 22041

8. PERFORMING ORGANIZATION REPORT NUMBER
--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-IKD
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARI Technical Report 1009

11. SUPPLEMENTARY NOTES
--

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

To determine the relationship between scores on a device-based test of gunnery proficiency and live-fire Tank Table VIII scores, a pooled sample of 73 Army National Guard (ARNG) tank crews (i.e., 24 from Smith and Hagman (1992) and 49 from this investigation) completed a 1-hour gunnery proficiency test on the Mobile Conduct-of-Fire Trainer (M-COFT) and then fired Table VIII the next day as part of annual training. For this pooled sample, a significant correlation was found between M-COFT test and Table VIII scores ($r = .67$, $p < .0001$). Based on the results of linear regression analyses, a tool was developed to predict Table VIII scores from M-COFT test performance measured in terms of hit rate or, easier to calculate, percentage of first-round kills. Although field tryouts are needed to verify the accuracy of the predictions, the results in this report suggest that device-based prediction of live-fire tank gunnery performance is possible, and that this prediction capability can be used by ARNG company commanders to assess the proficiency of tank crews and their need for additional training before live-fire gunnery evaluation on Table VIII.

14. SUBJECT TERMS
Reserve component
Tank gunnery
Training devices
M-COFT
Armor training

15. NUMBER OF PAGES
35

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

Technical Report 1009

Predicting Table VIII Tank Gunnery Performance From M-COFT Hit Rate

Monte D. Smith
CAE-Link Corporation

ɔseph D. Hagman
.S. Army Research Institute

Reserve Component Training Research Unit
Ruth H. Phelps, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

October 1994

Army Project Number
2Q263007A793

Human Factors in Training
Operational Effectiveness

Approved for public release; distribution is unlimited.

FOREWORD

The Army National Guard (ARNG) is emphasizing the use of training devices to enhance home-station training of M1 tank gunnery. To this end, work is under way to develop a device-based tank gunnery training and evaluation strategy for ARNG use at the company level. This report describes the results of research performed to (a) determine the relationship between tank crew gunnery scores on a device-based (i.e., Mobile Conduct-of-Fire Trainer [M-COFT]) test of gunnery proficiency and first-run live-fire scores on Tank Table VIII, and (b) develop a device-based tool for U.S. Army National Guard unit trainers to use in predicting individual crew qualification on Table VIII.

The research was conducted by the U.S. Army Research Institute for the Behavioral and Social Sciences, whose mission is to improve the effectiveness and efficiency of Reserve Component (RC) Training Research Unit, training through use of the latest in training technology. The research task support this mission, "Train Up: Technology-Based RC Training Strategies," is organized under Science and Technology Objective V.B.7 Unit Training Strategies.

The National Guard Bureau (NGB) sponsored this research under a Memorandum of Understanding signed 12 June 1985. Results have been presented to Chief, Training Division, NGB; Chief, Training Division, Office of the Chief, Army Reserve; and Special Assistant to the Commanding General, U.S. Army Armor Center.

EDGAR M. JOHNSON
Director

PREDICTING TABLE VIII TANK GUNNERY PERFORMANCE FROM M-COFT HIT RATE

EXECUTIVE SUMMARY

Requirement:

To (1) conduct a validation test of Smith and Hagman's device-based (i.e., Mobile Conduct of Fire Trainer [M-COFT]) predictions of first-run Table VIII tank gunnery scores, (2) construct a revised prediction model, if necessary, and (3) based on the new model, develop a practicable prediction tool for use at the company level in predicting the Table VIII scores of Army National Guard (ARNG) tank crews.

Procedure:

Forty-nine M1 tank crews from five companies of the 1-303rd Armor Regiment of the Washington ARNG and one company from the 3rd Battalion, 116th Cavalry Brigade of the Oregon ARNG served as participants. The tank commander and gunner from each crew took a 1-hour M-COFT test of gunnery proficiency and then fired Table VIII the next day as part of annual training.

Findings:

Data obtained from the present sample of 49 crews revealed that Smith and Hagman's M-COFT test-based predictions of Table VIII performance predictions were more accurate than expected from chance alone but of little practical value. To enable the development of a revised prediction model based on a larger sample of tank crews, the M-COFT test and Table VIII scores from the present sample were pooled with the analogous scores from Smith and Hagman's sample (n = 24). For this pooled sample of 73 crews, a significant correlation was found between M-COFT test and Table VIII scores (r = .67, p < .0001). Based on the results of linear regression analyses, a tool was developed to predict Table VIII scores from M-COFT test performance measured in terms of hit rate or, easier to calculate, percentage of first-round kills.

Utilization of Findings:

The findings confirm that a robust M-COFT-to-Table VIII performance relationship exists. The predictive tool developed on the basis of this relationship can be used by ARNG unit commanders to assess the proficiency of tank crews and their need for additional training prior to live-fire gunnery evaluation on Table VIII. Data from future field implementation tryouts are needed to verify the accuracy of the predictions. In the meantime, the findings of this research represent a step forward in the development of a device-based tool for predicting ARNG live-fire tank gunnery success.

PREDICTING TABLE VIII TANK GUNNERY PERFORMANCE FROM M-COFT HIT RATE

CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREDICTING TABLE VIII TANK GUNNERY PERFORMANCE FROM M-COFT HIT RATE

Background

Total Force Policy requires that the Army's Reserve Component (RC) soldiers attain and maintain readiness standards comparable to those of their Active Component counterparts. Because of constraints on time and limited access to range/maneuver areas, the majority of RC armor training must be accomplished at home station (i.e., armory or reserve center), where it is difficult to provide the kind of realistic tank gunnery training necessary to ensure skill proficiency.

Tank Gunnery Simulation

To increase RC home-station training capability (especially for combat arms units), the National Guard Bureau is seeking to use technology in the form of simulators and training devices. To guide the use of this technology and thereby promote the successful RC transition from equipment-based to device-based training in the area of tank gunnery, Morrison, Campshure and Doyle (1991) developed a strategy to link device-based training with on-tank performance. Under this strategy, the purpose of device-based training is to prepare individuals, crews, and platoons to be trained on the tank combat tables, with these tables providing the intermediate and terminal performance objectives for gunnery training.

The strategy has three phases: (1) begin with basic device-based training at home station, (2) proceed to intermediate device-based training at home station coupled with on-tank training at the Local Training Area (LTA), and (3) conclude with live-fire tank combat table evaluation at the Major Training Area (MTA).

Mobile Conduct-of-Fire Trainer (M-COFT)

The centerpiece device for this strategy is the M1 M-COFT, a computer-based tank gunnery simulator wherein tank commander (TC) and gunner pairs are placed in simulated crew stations and presented with a wide range of target engagement situations. The crew stations replicate interior features, dimensions, and lighting of the M1 tank, including weapons, sights/optics, and TC/gunner fire control systems. TC/gunner pairs follow actual engagement procedures, striving to produce "kills" of computer-generated target images. The M-COFT simulates all major TC/gunner M1 tank components across a variety of potential operating conditions (see Campshure, 1991, p. 12; pp. 21-22). Functionally equivalent to its predecessor, the U-COFT (Unit Conduct-of-Fire Trainer; U.S. Army Armor Center, 1985), the M-COFT can be moved from site to site because it is mounted on an

enclosed flatbed truck. The M-COFT's mobility reflects its intended purpose of fulfilling the unique training needs of armor units in the Army National Guard (ARNG). These units often have company-sized elements that are geographically dispersed from the rest of the battalion.

Out of six devices included in the Morrison, Campshure and Doyle (1991) strategy, M-COFT was allocated over one third of recommended basic and intermediate training hours. According to Morrison, Campshure, and Doyle (1991), and Morrison, Drucker and Campshure (1991), M-COFT simulates more gunnery training requirements than any other simulation device, and is the only device that can support the training of some gunnery related tasks (e.g., simultaneous engagements and most degraded mode gunnery procedures).

Relationship Between M-COFT and Table VIII

If the M-COFT device simulates a broad spectrum of M1 components and condition parameters, it is not unrealistic to expect that proficiency on the M-COFT should correlate with M1 tank gunnery scores. If crews can be trained to M-COFT proficiency at home station, and this training transfers to subsequent live-fire tank combat evaluations, numerous advantages will result, including reduced training costs, more efficient allocation of training time, and reduced ammunition requirements.

Theoretically, it should be possible to examine M-COFT scores and predict which crews will successfully qualify first-run on Table VIII. Support for this notion, however, is only beginning to accumulate. Morrison, et al. (1991) reviewed four investigations that reported either no significant correlations between M-COFT and live-fire gunnery performance (Butler, Reynolds, Kroh, & Thorne, 1982; Kuma & McConville, 1982; Hughes, Butler, Sterling, & Bergland, 1987) or a few low correlations on the speed, but not the accuracy, of performance (Black & Abel, 1987). One reason for this inconsistent relationship may be unreliable M-COFT gunnery scores. Graham (1986), however, reported significant M-COFT test-retest reliability coefficients, based on data samples of approximately 15 to 20 minutes at each administration. Six of his nine measures produced reliability coefficients of .70 or greater, and three of the six were in excess of .80. DuBois (1987) successfully replicated Graham's findings, although the obtained reliability coefficients were lower. Smith and Hagman (1992), moreover, reported statistically significant part-whole correlations (a form of internal consistency) between individual M-COFT exercises and a composite consisting of four exercises.

The first clear demonstration of an M-COFT-to-live-fire gunnery relationship suggested the critical importance of a methodological factor. Campshure and Drucker (1990) reported

significant bivariate correlations between Table VIII total score and either crew Reticle Aim Level or TC Reticle Aim Level of the M-COFT training matrix. Campshure and Drucker hypothesized that a composite measure of M-COFT achievement (e.g., M-COFT matrix position, based on aggregated sessions) may provide a more reliable prediction of Table VIII performance than scores from a few M-COFT engagements.

Both empirical and logical grounds support Campshure and Drucker's (1990) contention, and suggest that composite measures are more stable (and hence more reliable) than specific test performance scores. Table VIII performance represents a multi-faceted composite of many behaviors (including cognitive, motivational, and perceptual-motor functioning) as well as quality, extent and intensity of prior training. Because of the complexity of the criterion measure, only a composite sampling of M-COFT performance, encompassing a broad array of specific M-COFT behaviors, can reasonably be expected to predict Table VIII outcomes. It is a psychometric axiom that longer tests are more reliable, and a test has to be reliable to be valid. Reliability, in fact, sets an upper limit to validity (Black & Graham, 1987). Campshure and Drucker (1990) used composite measures on both sides of their prediction equation. Table VIII total score (the sum of 10 engagements, or "tasks") served as the live-fire composite measure, and M-COFT matrix position (a broad but admittedly undifferentiated aggregate) served as their primary composite predictor.

Prior to Campshure and Drucker's (1990) composite methodological approach, much of the previous research into the M-COFT-to-live-fire relationship was characterized by limited data samples on both the predictor and criterion sides of the prediction equation. That is, although numerous tank crews were usually observed, only a limited sample of data was collected from each crew, sometimes as little as one M-COFT exercise and a single live-five gunnery exercise. A typical procedure involved calculation of bivariate correlations between a single M-COFT exercise and a single live-fire gunnery exercise (Butler, et al., 1982; Kuma & McConville, 1982), or discrete speed and/or accuracy measures and live-fire performance (Black & Abel, 1987; Hughes, et al., 1987).

Smith and Hagman (1992) investigated the relative utility of composite versus more discrete measures of M-COFT achievement for predicting Table VIII scores, and concluded that a composite measure is superior. A composite measure of M-COFT achievement (Hoffman & Witmer's [1989] Hit Rate), based on four combined M-COFT exercises, produced a significant correlation with Table VIII scores ($r[24] = .65$, $p < .01$). This was not the case when Hit Rate was based on individual exercises. Individually, three of the four exercises produced nonsignificant correlations ranging from .21 to .27, and Hit Rate based on the fourth and

most predictive exercise, although significantly correlated with Table VIII, accounted for only 18% of the Table VIII variance, versus 42% that was accounted for by the four-exercise composite. The best combination of three exercises accounted for only 26% of Table VIII variance. Smith and Hagman concluded that short M-COFT tests appear to place a severe limitation on the ability to demonstrate a gunnery training device's correlation with Table VIII criterion measures.

Addition of demographic variables enhanced the power of the Smith and Hagman (1992) prediction equation. Using their full multiple regression model, based on three predictors, they algebraically solved for the M-COFT Hit Rate necessary in order to predict a mean Table VIII score of 700 (the score required for qualification), as well as the necessary M-COFT Hit Rate for predicting (with 95% confidence) a minimum Table VIII score of 700.

Implications of Transitioning to M-COFT Training

Despite the central importance of M-COFT in the proposed transition to device-based RC tank gunnery training, little is known about how the device should be used to best facilitate subsequent Table VIII qualification. More practical guidance is needed for setting M-COFT proficiency objectives. Although the Smith and Hagman (1992) model holds promise as an aid for guiding RC device-based tank gunnery training efforts, it also embodies several limitations, including a small sample size upon which it was based, a low mean Table VIII score in the original database, and inherent complexity due to its multivariate nature (i.e., the use of both M-COFT scores and demographic variables as predictors). For these reasons their model must be cross-validated on different samples of armor crews before it can be used to guide device-based tank gunnery training. If their reported relationship between M-COFT Hit Rate and Table VIII scores is cross-validated, the replication data can be combined with the original data to produce an improved prediction model.

In addition to psychometric limitations, a practical limitation of the Smith and Hagman (1992) model concerns their inclusion of demographic variables in the prediction equation. Although demographics boosted the strength of the equation, it also made the model more complex and hence less practicable from the standpoint of RC armory training implementation. From an implementation standpoint, the ideal model would provide unit commanders with an exclusively M-COFT-based yardstick whereby crews could be trained to levels of device proficiency corresponding to known probabilities of subsequent Table VIII qualification. For example, a given level of M-COFT proficiency might be associated with a 70% probability of first-run Table VIII qualification, a slightly higher proficiency with 80%, and

so on, up to confidence levels stringent enough to satisfy the most exacting training standards.

Even with a simplified yardstick based exclusively on M-COFT measures, RC unit commanders would still be faced with the challenge of deriving Hoffman and Witmer (1989) Hit Rate scores, a task that is both computationally complex and tedious. For optimal use of the M-COFT training device, software modifications are needed that would, at the punch of a console button, compute a crew's Hit Rate score at the conclusion of a training session and compare and contrast it with scores obtained in earlier training sessions. This, however, would be a major software programming modification. An alternative solution might involve development of a personal computer (PC) algorithm to permit M-COFT instructor/operators (I/Os) to punch in key measures following a training session and have the computer calculate Hit Rate. Even this procedure, however, would saddle the I/O with considerable interpretative responsibilities and require the transfer of several dozen measures to a PC. The simplest interim solution would entail identification of the best predictors of Hit Rate, perhaps one or two measures, that could be obtained quickly from current M-COFT printouts and then easily converted to estimates of Hit Rate.

Purpose

The present research: (1) focused on the predictive utility of the key component from Smith and Hagman's (1992) Table VIII prediction model: Hoffman and Witmer's (1989) Hit Rate score, (2) provided a validation test of the Smith and Hagman (1992) Hit Rate prediction model, based on a different and larger sample of RC tank crews, with strict adherence to standardized M-COFT test administration procedures used in the original research, (3) provided an opportunity to pool M-COFT and Table VIII scores from the validation sample with corresponding measures from the original sample to develop a revised M-COFT-based Table VIII prediction model, (4) permitted development of M-COFT-based cut-scores, using pooled data, for setting M-COFT proficiency standards, and (5) identified the best and most easily derivable (from M-COFT printouts) surrogate measures for Hoffman and Witmer's (1989) Hit Rate.

Method

Participants

Forty-nine M1 tank crews from five companies of the 1-303rd Armor Regiment of the Washington ARNG and one company from the 3rd Battalion, 116th Cavalry Brigade of the Oregon ARNG served as participants.

General Procedure

Crews took the M-COFT Test of Gunnery Proficiency (CTGP) and then fired Table VIII the next day as part of Annual Training (AT), 1993. Data from Washington crews were collected at the Multi-Purpose Range Complex (MPRC) at the Yakima Training Center. Data from the Oregon crews were collected from the MPRC at Orchard Range near Gowen Field (Idaho).

The M-COFT Test of Gunnery Proficiency (CTGP)

Hoffman and Witmer's (1989) CTGP provided the composite assessment of M-COFT gunnery skills. The CTGP consists of four M-COFT matrix exercises (duration of administration approximately 1 hr) selected to correspond to conditions that occur in Table VIII. The exercises cover: target arrays, ranges, own tank movement, target movement, Nuclear, Biological and Chemical (NBC), crew configuration (four- or three-man), day/night, and number of targets per engagement. The selected M-COFT exercises do not replicate Table VIII tasks exactly, but they represent all Table VIII conditions in somewhat different sequences and combinations. Figure 1 lists exercises included in the CTGP, with major Table VIII conditions represented. Hoffman and Witmer (1989) provide extended discussion on the rationale for selecting M-COFT exercises and the overlap among conditions represented in Table VIII, the CTGP, and the known domain of M1 gunnery conditions.

Standardized Administration. A standardized set of CTGP administration procedures (Hoffman & Witmer, 1989) was followed closely, including verbatim reading of instructions at the beginning of each exercise. The test procedures emphasized testing requirements, rather than usual M-COFT training needs. For example, no feedback or coaching was provided during testing, and switch setting instructions were given only at the start of each exercise.

(I/O) Training. To help ensure consistency of CTGP administration, the number of I/Os was limited to six, and the principal I/O was the same master gunner as in the Smith and Hagman (1992) investigation. (Multiple I/Os were required in order to avoid fatigue from consecutive testing sessions.) The principal I/O, in both the present and original investigation, administered approximately two-thirds of all CTGP protocols. All I/Os were provided an overview of research objectives through briefings that highlighted differences between CTGP testing and normal M-COFT training. Briefings also emphasized the critical importance of withholding feedback. I/Os practiced the test administration scenario prior to their first actual testing session, and observed each other's performance during early sessions to ensure that the standardized set of administration procedures was being followed. Either a principal researcher or

EXERCISE	CONDITIONS REPRESENTED
1	Stationary own tank Multiple moving and stationary targets Daylight with unlimited visibility Simultaneous engagement Modified to include NBC conditions Modified to cover three-man crew engagement
2	Moving own tank Multiple moving and stationary targets Tank, helicopter, troop and armored personnel carrier (APC) targets Battlefield conditions Friendly M1 tank in one engagement Firing from other vehicles depicted in the scene Visibility reduced by fog
3	Night gunnery Stationary own tank Multiple, moving and stationary targets Tank, APC, and helicopter targets Modified to include NBC conditions Modified to cover three-man crew engagement
4	Stationary and moving targets Gunner's auxiliary sight (GAS) Firing from short halt

Figure 1. Conditions represented in COFT Test of Gunnery Proficiency (CTGP) exercises (Exercise 1 = M-COFT Exercise 34611; Exercise 2 = 34633; Exercise 3 =34622; Exercise 4 = 31563).

the principal I/O was present during test administrations in case questions or problems arose.

Hit Rate. The CTGP produces a composite measure of gunnery proficiency -- a "test-wide" M-COFT performance measure that is weighted for the number of targets in each of 22 contributing engagements. This composite measure of gunnery proficiency is called Hit Rate, which Hoffman and Witmer (1989) define as:

$$\underset{\text{(hits/time)}}{\text{Hit Rate}} = \underset{\text{(hits/rounds)}}{\text{Hit Proportion}} \times \underset{\text{(rounds/time)}}{\text{Fire Rate}}$$

"Hit rate, adjusted for hits on friendly targets, is the recommended metric for assessment of overall crew proficiency.

Hit rate is calculated for each engagement from information on M-COFT printouts on rounds fired, kills, and time. Overall hit rate is calculated from the weighted averages for firing rate and hit probability, where engagement firing rates and hit probabilities are weighted by the number of targets in the engagement." (Hoffman & Witmer, 1989; p 29)

Hit Rate, as defined by Hoffman and Witmer (1989) is a metric that incorporates hits, rounds fired, and time. The resulting metric is weighted by the number of available targets. The unit of time used by Hoffman and Witmer (1989), and by present investigators, was seconds. This time unit, however, is arbitrary, and other time units could be substituted.

Although the scoring procedure for Hit Rate is computationally complex and laborious, it does capture in a single metric the essential elements of gunnery success: rounds fired, time expended, accuracy of fire, and completeness (were all threat targets hit?). (For more details on the Hit Rate scoring procedure, refer to Hoffman and Witmer, 1989.) In addition to the overall Hit Rate which is based on a weighted combination of all exercises, similar scores can be calculated for individual exercises. Additionally, standard M-COFT computer printouts provide numerous subsidiary measures, all of which can be compared and contrasted with the predictive utility of the various Hit Rate measures.

Table VIII

Each crew fired six day engagements and four night engagements, selected from among twelve engagements described in FM 17-12-1 (Department of the Army, 1988). Crews received raw scores of from 0 to 100 on each engagement, based on engagement speed, accuracy of fire, and threat capability. Penalty points (crew cuts) were deducted from each engagement raw score based on observed procedural errors. Scores were summed (after deduction of penalty points) for the six day engagements, producing a total day score. Similarly, scores were summed for the four night engagements, producing a total night score. Day score and night score were summed to produce a total Table VIII score for each crew. Total scores could range from 0 to 1,000, and a score of 700 was required for crew qualification.

Results

Relationship Between M-COFT Hit Rate and Table VIII Scores

Table VIII scores. Table VIII scores ranged from 141 to 913, with a mean of 563 and a standard deviation of 147. Seven of the 49 crews (14.3%) qualified on their first-run, obtaining scores above 700. Table VIII scores were significantly higher ($F[1,71] = 27.1$; $p < .0001$) than those reported by Smith and Hagman

(1992). In that investigation, the range was from 59 to 703, with a mean of 356, a standard deviation of 184, and a 4.2% first-run qualification rate.

M-COFT Hit Rate. Hit Rate scores ranged from .015 to .062, with a mean of .034 and a standard deviation of .012. Hit Rate scores were significantly higher ($F[1,71] = 31.7$; $p < .0001$) than those reported by Smith and Hagman (1992). In that investigation, the range was from .005 to .032, with a mean of .018 and a standard deviation of .007. Higher hit rate scores are consistent with more hours of M-COFT practice. In the earlier investigation, Smith and Hagman reported that because M-COFT devices had arrived at armories only days before the research began, crews had received an average of only 2 hr prior familiarization with the device. Crews in the present investigation had received substantially more M-COFT training, averaging 68 hr for TCs and 43 hr for gunners.

Correlation between M-COFT Hit Rate and Table VIII. Five Hit Rate scores were calculated, the first four were based on individual M-COFT exercises and the fifth was based on the combination of all four exercises. Correlations between these measures and Table VIII total scores are summarized in Table 1, for both Smith and Hagman data (1992) and for data from the present investigation.

Table 1

Correlations Between Mobile Conduct-of-Fire Trainer (M-COFT) Hit Rate and Table VIII Total Score for Smith and Hagman (1992) and for the Present Investigation

(Number of Tank Crews)	Exercise 1	Exercise 2	Exercise 3	Exercise 4	Composite 1-4
S&H (24)	$.42^{a}$	.21	.27	.20	$.65^{b}$
Present (49)	$.37^{b}$	$.49^{b}$	.28	.22	$.53^{b}$

Note. Numbers in parentheses refer to the number of tank crews in the two investigations.
^{a} $p < .05$
^{b} $p < .01$

In the present investigation, Hit Rates based on two individual exercises correlated significantly with Table VIII scores. Exercise 1 (Exercise Number 34611 in the M-COFT training matrix) correlated significantly with Table VIII in both investigations. Exercise 2 (Number 346333 in the M-COFT training

matrix) correlated significantly in the present investigation, but not in the earlier investigation. In both investigations, Hit Rate based on a composite of all four exercises produced a higher correlation with Table VIII than any individual exercise, although the difference between the composite and the most predictive individual exercise was not as great in the present investigation as it was in the earlier one. Note, however, that the most predictive individual exercise in the present investigation produced a nonsignificant correlation in the earlier investigation. Only Hit Rates based on Exercise 1 engagements consistently predicted Table VIII scores across the two investigations. In the earlier investigation, Exercise 1 Hit Rate accounted for 42% as much variance as composite Hit Rate ($.42^2/.65^2 = .42$). In the present investigation Exercise 1 Hit Rate accounted for 49% as much variance as composite Hit Rate ($.37^2/.52^2 = .49$).

Consistent with Smith and Hagman (1992), composite Hit Rate emerged as the superior predictor of Table VIII scores. In both investigations, composite Hit Rate accounted for a substantial proportion of Table VIII score variance.

Validating an M-COFT-based Table VIII Prediction Model

Based on data from the Smith and Hagman (1992) report, it is possible to develop a prediction model and then test the validity of that model with data from the present investigation. Data from the earlier investigation were used to develop a linear regression prediction equation of the form:

$$Y = B_0 + B_1(X_1)$$

where Y is the predicted Table VIII score, B_0 is the intercept (or theoretical Table VIII score when M-COFT Hit Rate is set equal to zero), B_1 is the empirically determined regression coefficient that links changes in Table VIII scores with changes in M-COFT Hit Rate, and (X_1) is M-COFT Hit Rate. The following equation resulted, with a standard error of estimate of 143.8:

$$Y = 42.59 + 17{,}056(\text{Hit Rate})$$

Following the procedure of Smith and Hagman (1992), we can set Y = the minimum Table VIII qualifying score and algebraically solve for the single unknown: the M-COFT Hit Rate necessary in order to predict a Table VIII score of 700. Using the regression equation based on 24 crews from the Smith and Hagman (1992) investigation (with Multiple R = .65 and $R^2 = .42$), we have:

$$700 = 42.59 + 17{,}056(\text{Hit Rate})$$

Which reduces to:

$$657.41 = 17{,}056(x)$$

Solving for the unknown (Hit Rate) we have:

$$17.056(x) = 657.41$$

$$x = .039$$

Therefore, the model predicts that a new tank crew trained to an M-COFT proficiency as indicated by a Hit Rate score of .039, will shoot a Table VIII score of 700. However, the predicted Table VIII score of 700 is a point estimate, and a new tank crew's actual Table VIII score may be somewhat greater or somewhat lesser than this predicted score. An individual tank crew has an equal probability (50%) of scoring either above or below the predicted score of 700. A more pertinent issue is: What M-COFT Hit Rate is necessary in order to predict, with a specified level of confidence, that a new crew will qualify on Table VIII? This is a tricky issue for several reasons. To paraphrase Hays (1963):

[Be sure to note the interesting fact that the regression equation found for a sample is not equally good ... over all the different values of the predictor variable. The prediction is at its best when the predicted score is the same as the mean predicted score, since the confidence interval is smallest at this point. However, as predicted values grow increasingly deviant from the mean predicted score (in either direction) the confidence intervals grow wider. For the more extreme values of the predicted score, we can have little confidence that the actual mean obtained for a sample of individuals (each showing the same predicted score) will be anywhere near what we have predicted.]

Based on data from the earlier investigation (Smith & Hagman, 1992), predicting the level of M-COFT Hit Rate that will ensure a minimum Table VIII score of 700 takes us beyond the observed range of data values. Statistical extrapolation beyond the range of observed values is always risky. That the relationship between predictor and criterion variables will remain essentially unchanged outside the bounds of observed data values is an unsubstantiated assumption. For these reasons, the following analyses must be considered as speculative.

Nevertheless, by regressing Hit Rate scores upon Table VIII scores, we obtain a standard error of the estimate with which we can construct a (one-tailed) confidence interval on the upper side of the desired mean criterion score of 700. The upper end of this one-tailed confidence interval can then be plugged into the prediction model to identify the necessary level of M-COFT Hit Rate that must be achieved in order to predict with 85% confidence that a minimum Table VIII score of 700 will be obtained.

The upper bound of the 85% (one-tailed) confidence interval is 879 (see Hays [1963], p. 523, equation 15.22.4) for computational formulae). If we wish to ensure that 85% of crews shoot at least 700 on the next Table VIII exercise, and we assume that confidence intervals outside the bounds of observed values will be similar to those on the fringe of observed values, we can calculate the M-COFT Hit Rate necessary to predict (with 85% confidence) a minimum mean score of 700 on the Table VIII exercise by plugging the target value of 879 into the prediction equation derived above, as follows:

$$879 = 42.59 + 17{,}056(\text{Hit Rate})$$

Which reduces to:

$$836.41 = 17{,}056(x)$$

Solving for the unknown (Hit Rate) we have:

$$17{,}056(x) = 836.41$$

$$x = .049$$

To test the predictive utility of these two Hit Rate cut-scores, two validation tests were conducted, the first to test the predictive efficacy of the .039 Hit Rate cut-score, and the second to test the .049 Hit Rate cut-score. It will be recalled that the first cut-score, .039, was the estimated Hit Rate required to produce a Table VIII score of 700. The second cut-score, .049, was the estimated Hit Rate required to produce a Table VIII score of at least 700. The validation test consisted of examining the data records of each of the 49 crews in the present investigation to determine whether they met or surpassed the required Hit Rate criteria derived from the earlier sample, and whether or not they obtained first-run qualification by shooting a Table VIII score of 700 or greater. Table 2 presents the outcomes of these two tests.

Using the cut-score of .039, the model predicted that 16 crews would qualify (the first row total in the top half of Table 2). Only 7 crews, however, actually achieved first-run qualification (the first column total in the top half of the table). This suggests that the predicted Hit Rate of .039, based upon $\underline{n} = 24$ crews from the earlier investigation, was too low to ensure Table VIII qualification. Based on this cut-score, over twice as many crews were predicted to qualify as actually qualified. Of the 16 crews predicted to qualify, 4 actually achieved first-run qualification, for a prediction accuracy of 25% (4/16). Although the .039 cut-score obviously over-predicts qualification (16 predicted versus 7 qualified), it does add some predictive accuracy above what would be known without any model. From the marginal qualification totals alone, we know only that 7 of 49

Table 2

Validation Tests for Two Levels Of A Table VIII Prediction Model (Cell Values Represent Numbers of Crews)

Cut Score = .039

		Qualified Yes	Qualified No	Total
Predicted	Yes	4	12	16
	No	3	30	33
	Total	7	42	49

Cut Score = .048

		Qualified Yes	Qualified No	Total
Predicted	Yes	3	4	7
	No	4	38	42
	Total	7	42	49

crews (14.3%) successfully qualified. Without additional information, attempting to identify individual crews that qualified would be a random process, and would yield a success rate of 14.3%. A chi square test (1, $\underline{n}$ = 49) produced a non-significant value of 2.22, indicating that for this sample, 25% accuracy of prediction does not differ, statistically, from 14.3%. Thus, actual qualification was statistically independent of predicted qualification and the model was ineffective. With a correction for continuity, this test value reduced to 1.12.

A statistical test more precisely fitted to the hypothesis of concern in this instance is the relative incidence ratio, a variant of the relative risk ratio. This test measures the strength of association between presence or absence of a factor (predicted to qualify versus not predicted to qualify in this instance) and incidence of the predicted outcome (qualification versus non-qualification). The test computes a ratio between incidence of qualification given prediction and incidence of qualification given non-prediction. In this case, 4 of 16 crews

(25%) qualified given prediction and 3 of 33 crews (9.1%) qualified given non-prediction. The ratio of these two incidence rates (25.0/9.1 = 2.75) is not statistically significant.

In the bottom half of Table 3, using the cut-score of .049, the model predicted that 7 crews would qualify (the first row total). This predicted number matches the number of crews that actually qualified on the first run (n = 7), suggesting that .049 may be closer to a useful Hit Rate cut-score. Of the 7 crews predicted to qualify based on having obtained a Hit Rate score of at least .049, 3 crews actually achieved first-run qualification, for a prediction accuracy of 42.9% (3/7). Applying the same tests as before, a chi square (1, n = 49) yielded a value = 5.44, $p < .05$, which was reduced to a non-significant 3.06 by a correction for continuity adjustment. The relative incidence test produced a ratio of 4.50 (the incidence of qualification given prediction [3/7 = 42.9%] divided by the incidence of qualification given non-prediction [4/42 = 9.5%]), which is significant at $p < .05$. Because the chi square test yielded borderline results and the relative incidence ratio indicated statistical significance, a Fisher's exact test also was calculated. It produced an outcome with an associated probability = .05.

The Hit Rate cut-score of .039 was clearly too low, as indicated by nonsignificant test values and by the fact that it predicted twice the number of qualifying crews as actually qualified (16 predicted versus 7 qualified). The .049 cut-score was more successful. It accurately predicted the number of successfully qualifying crews, and its accuracy of prediction (42.9%) was three times as effective as what could be obtained (14.3%) without benefit of any prediction model. However, tests of association indicated that the model was only marginally successful, statistically.

The marginal success of the prediction model may be partly attributable to the fact that one of the predicted scores (700) was on the upper fringe of Table VIII values that were observed in the earlier investigation (Smith & Hagman, 1992) upon which the prediction model was based, and the other predicted score (879) was completely beyond the observed range of data values. As cautioned above, statistical extrapolation beyond the range of observed values is always risky. Such extrapolation assumes that the relationship between predictor and criterion variables will remain essentially unchanged outside the bounds of observed data values. In this instance, that assumption may have attenuated the predictive accuracy of the prediction model. The reader will recall that the mean Table VIII score (356) for the 24 crews upon which the prediction model was based was over 200 points less than the mean Table VIII score (Mean = 563, n = 49) obtained by crews in the present investigation. This suggests that a prediction model based on data from the present investigation

would produce Hit Rate cut-scores higher than the ones produced by a model based upon the earlier data.

To test this hypothesis, the Hit Rate scores required for predicting a Table VIII score of 700 and a Table VIII score of at least 700 (with 85% confidence) were calculated, using the steps outlined above for the earlier sample, based on $\underline{n} = 49$ crews from the present investigation. The results, along with cut-scores based upon the earlier ($\underline{n} = 24$) investigation, are summarized in Table 3, which shows that estimates of required M-COFT proficiency based on the earlier sample were conservative. That is, required Hit Rate estimates based on the present sample are higher than the required Hit Rate estimates based on the earlier sample.

Table 3

Mobile Conduct-of-Fire Trainer Hit Rate Needed in Order to Predict Mean and Minimum Table VIII Scores of 700

		M-COFT Hit Rate Needed to Predict	
Investigation	$\underline{n}$	A Score of 700	At Least 700
S&H (1992)	24	.039	.049
Present	49	.055	.074

Note. S&H = Smith and Hagman

A Prediction Model Based on Pooled Data

When data from the two investigations were pooled, the Table VIII mean was 495, with a standard deviation of 187. The composite Hit Rate mean was .029, with a standard deviation of .013. The correlation between Table VIII scores and composite Hit Rate was $\underline{r}(73) = .67$; $\underline{p} < .0001$. This coefficient of correlation based on pooled samples was greater than that obtained from either sample separately. (The coefficient was .65 in the earlier sample and .53 in the present investigation.) Figure 2 shows how the two distributions of scores combined to produce a substantial coefficient of correlation, notwithstanding the fact that mean scores for the two samples differed significantly on both variables. A datum from the present investigation appears in Figure 1 as a capital N, and a datum from the earlier investigation is represented as a capital O. Although mean scores for the two groups differed significantly (on both variables), Figure 2 shows that both distributions

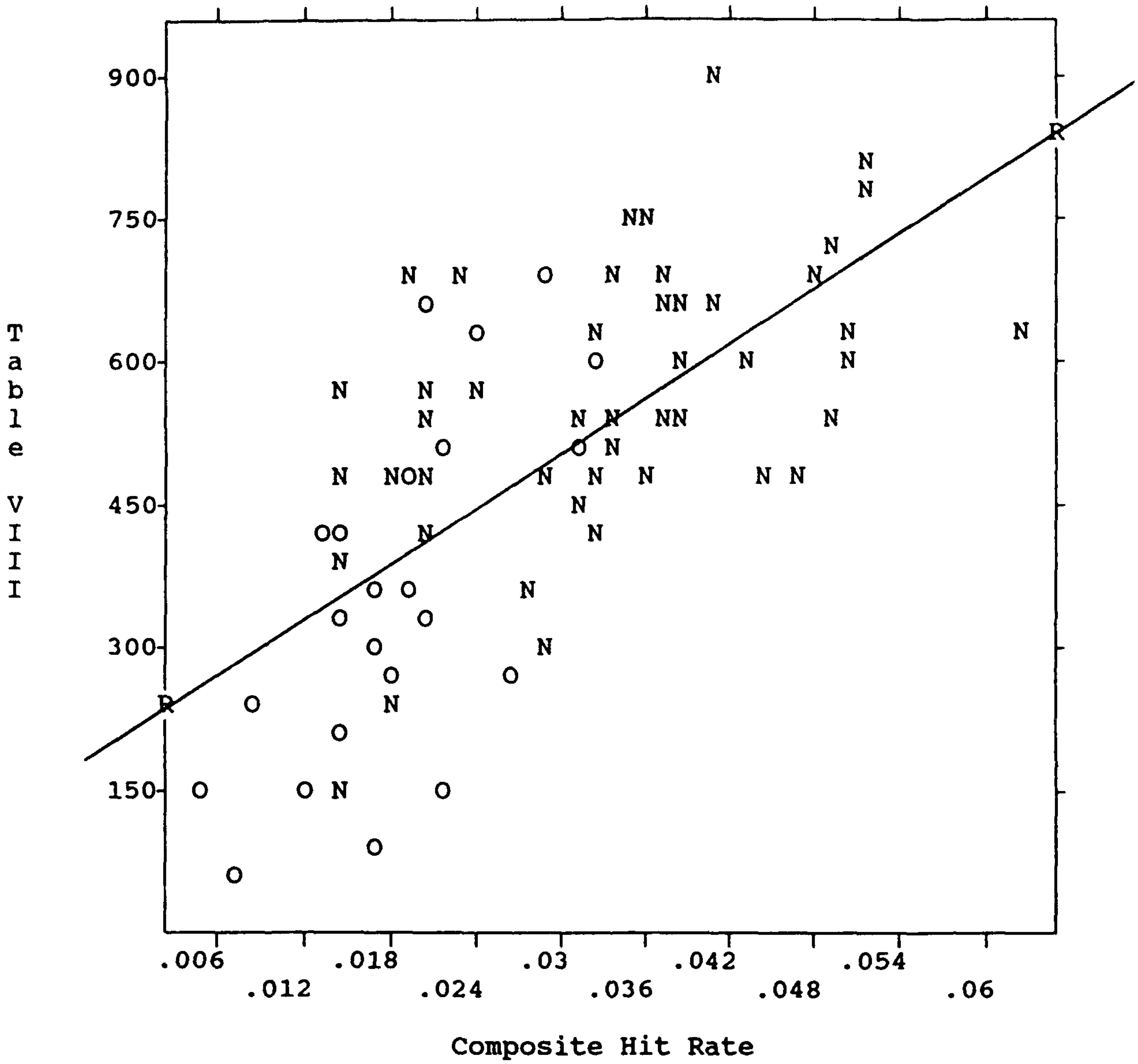

Figure 2. Plot of M-COFT Composite hit rate with Table VIII total score (*N* = 73; *r* = .67; *p* < .0001). N = "New" data from the present investigation and O = "Old" data from the earlier investigation. The "Rs" printed on the vertical axes indicate the points of intercept for the best-fit linear regression line.

combine to form a single linear continuum, with combinations of lower Hit Rate and lower Table VIII scores (from the earlier investigation) tending to cluster in the lower left quadrant, and combinations of higher Hit Rate and higher Table VIII scores (from the present investigation) tending to cluster in the upper right quadrant.

Although the pattern in Figure 2 suggests that both the earlier sample and the present one are from a single linear continuum, it also suggests that predictive models based on either sample will have limited predictive success because score dispersions in both samples are restricted. Because of the wide dispersion of both Table VIII scores and Hit Rate scores, and the sizeable correlation between the two variables in the pooled data, the pattern in Figure 2 suggests that a regression model based on the pooled data would yield greater predictive accuracy than a model based on either sample separately. The model for predicting Table VIII scores, using pooled data, was:

$$Y = B_0 + B_1(X_1)$$

Which yielded the following equation, with a standard error of estimate of 139.8:

$$Y = 214.88 + 9{,}795(\text{Hit Rate})$$

Following the procedure outlined above, we can set Y = the minimum Table VIII qualifying score and algebraically solve for the single unknown: the M-COFT Hit Rate necessary in order to predict a Table VIII score of 700:

$$700 = 214.88 + 9{,}795(\text{Hit Rate})$$

Which reduces to:

$$485.12 = 9{,}795(x)$$

Solving for the unknown (Hit Rate) we have:

$$9{,}795(x) = 485.12$$

$$x = .050$$

In the same manner as outlined above for the earlier sample of $\underline{n} = 24$, we can calculate the upper bound of the 85% (one-tailed) confidence interval (849) and plug it into the formula in order to determine the Hit Rate value necessary to ensure that 85% of crews shoot at least 700 on the next Table VIII exercise.

$$849 = 214.88 + 9{,}795(\text{Hit Rate})$$

Which reduces to:

$$634.12 = 9{,}795(x)$$

Solving for the unknown (Hit Rate) we have:

$$9{,}795(x) = 634.12$$

$$x = .065$$

Comparing these two cut-scores, based on pooled data, with those displayed in Table 3 for the two samples separately, indicates that the pooled cut-scores are of intermediate magnitude. Unfortunately, it is not possible to validate the pooled predictive model (or the model based on data from the present sample). That would require collection of another set of data, because a model cannot be validated on the same data upon which it is based. However, we can construct confidence bands around the predicted Table VIII score of 700, based on all 73 crews in the combined sample, and use this information to estimate the liklihood of a crew shooting at least 700.

Table 4 shows several confidence bands around the predicted score of 700. If a crew is trained to a level of COFT Hit Rate proficiency as indicated in the first column, its predicted Table VIII score can be determined by reading across to

Table 4

Probability Bands Around a Predicted Table VIII Score of 700, Expressed as Variations in the Predictor Variable, Hit Rate. Based on *N* = 73 Crews.

Hit Rate	Predicted Table VIII	Probability of Shooting A Table VIII of At Least 700
.031	517	10%
.037	578	20%
.042	626	30%
.046	664	40%
.050	700	50%
.053	736	60%
.057	774	70%
.062	822	80%
.065	849	85%
.068	883	80%
.074	935	95%

the second column. The third column is the confidence we may have that the crew will score at least 700 on Table VIII, given the Hit Rate score specified in the first column. For instance, the predicted Table VIII score for a crew that trains to a Hit Rate proficiency of .050 is 700. Not all such crews, however, will shoot exactly 700. Half will obtain a score higher than 700; half will obtain a score lower than 700. Hence, we can have only 50% confidence that this crew will obtain a Table VIII score of at least 700. For a crew trained to a Hit Rate level of .065,

however, we can predict a Table VIII score of 849, and we can have 85% confidence that the Table VIII score obtained by this crew with be at least 700. With lower Hit Rate scores, for instance .037, we would predict (based on the prediction model explained above) a Table VIII score of 578. However, there would also be a 20% possibility that this crew would obtain a Table VIII score of at least 700.

A Short-Cut Predictor of M-COFT Hit Rate

The results above indicate that Hit Rate has considerable utility as a predictor of subsequent Table VIII scores. Hit Rate, adjusted for hits on friendly targets, is based on rounds fired, hits, and time, and is calculated from weighted averages for firing rate and hit probability. Weights are determined by the number of targets in each of 22 engagements. Although Hit Rate does capture in a single metric the essential elements of gunnery success: rounds fired, time expended, accuracy of fire and completeness, the scoring procedure is computationally complex and laborious. However, standard M-COFT computer printouts also provide numerous subsidiary measures, one of which may bear a sufficiently strong relationship to Hit Rate in order to qualify as a substitute. An examination of M-COFT printouts for the pooled data ($N = 73$) revealed that the variable with the highest correlation with Hit Rate was proportion of first round kills ($r = .80$, $p < .0001$), defined as the number of targets killed with a first round divided by the total number of available targets in 22 engagements. Proportion of first round kills also correlated significantly with Table VIII scores ($r = .56$, $p < .0001$) This measure of device performance can be extracted readily by I/Os from M-COFT printouts at the conclusion of any training session that consists of at least one completed exercise.

When Table VIII scores and proportion of first round kill measures were entered into a regression routine, the following prediction equation resulted:

$$Y = 186.05 + 835.89(X)$$

where Y = predicted Table VIII and X = proportion of first round kills. The standard error of estimate was 156.19, based on a predictor mean of .370 and standard deviation of .124. The regression algorithm yielded an $F(1,71) = 31.68$, $p < .0001$, with an $R = .56$ and $R^2 = .31$.

With this prediction equation we can estimate the required proportion of first round kills that are required in order to predict a Table VIII score of 700, by plugging 700 into the above equation and solving for the only remaining unknown, proportion of first round kills:

$$700 = 186.05 + 835.89(x)$$

$$513.95+ 835.89(x)$$

$$x = .61$$

Seven hundred is the predicted Table VIII score for crews with .61 proportion of first round kills. However, only a few crews with a .61 proportion of first round kills will actually obtain a Table VIII score of 700. Most crews will be either above or below the predicted point estimate. For an individual crew, we can have only 50% confidence that its actual Table VIII score will be at least 700. Table 5 presents several confidence levels around the predicted score of 700. For example, if a crew obtains .82 proportion of first round kills, our best estimate of Table VIII is 870, and we may be 85% confident that its Table VIII score will be at least 700. For a crew with a comparatively low .36 proportion of first round kills, we can have only 10% confidence of its obtaining a Table VIII score of at least 700.

Table 5

Probability Bands Around a Predicted Table VIII Score of 700, Expressed as Variations in the Predictor Variable, Proportion of First-Round Kills. Based on N = 73 Crews.

pFRKills	Estimated Table VIII	Probability of Shooting a Table VIII of At Least 700
.36	491	10%
.45	561	20%
.51	615	30%
.57	659	40%
.61	700	50%
.66	741	60%
.72	785	70%
.78	839	80%
.82	870	85%
.86	909	80%
.93	967	95%

<u>Note</u>. pFRKills = proportion of first round kills

Discussion

This investigation replicated Smith and Hagman's (1992) finding of a significant positive linear relationship between a composite measure of M-COFT proficiency (Hit Rate) and Table VIII scores. The strength of the linear relationship in the present investigation (r = .53) was not as strong as in the earlier investigation (r = .65), but a comparative plot of Hit Rate and Table VIII scores from the two investigations suggested an explanation. Table VIII score distributions in both investigations only partially represented the total range of possible Table VIII and Hit Rate outcomes. In the earlier investigation, Table VIII and Hit Rate scores clustered on the low end of the potential score continuum. In the present investigation, both Table VIII and Hit Rate scores were significantly higher, and clustered in the middle of potential score outcomes. When the two distributions of Table VIII and Hit Rate scores were pooled, they produced a stronger linear relationship than when examined separately. Visual inspection of the pooled data suggested that the two samples covered approximately two-thirds of the potential Table VIII score continuum, and approximately the same proportion of potential Hit Rate outcomes. If a third sample were available to add to the pool, with a Table VIII mean of approximately 750, it seems reasonable to expect that the resulting relationship between Table VIII and Hit Rate might increase even further in strength.

The phenomenon of limited range is a methodological artifact known as truncation. It is axiomatic in measurement theory that truncation of range will either attenuate or eliminate observed relationships. That is, even if two variables are perfectly correlated (r = 1.00) in the population, samples drawn from that population that fail to capture an adequate range of available scores may produce attenuated correlations, and in extreme instances may produce no correlation. Figure 2 suggests that truncation of range may have suppressed the correlation (r = .53) between Hit Rate and Table VIII scores in the present investigation. This impact could also have occurred to some extent in Smith and Hagman (1992). (Possibly to a smaller degree because Table VIII variance was greater in the earlier investigation.) Indeed, part of the M-COFT-to-live fire relationship that occurred in the present research and in Smith and Hagman (1992) may be attributable to the substantial dispersion in Table VIII scores that occurred in both investigations. Smith and Hagman reported a range of 644 and a standard deviation of 184. In the present investigation, scores ranged from 141 to 913, with a standard deviation of 147. It is interesting to note that the only previous research to demonstrate a convincing M-COFT-to-Table VIII relationship (Campshure & Drucker, 1990) also reported a substantial Table VIII score dispersion (Range = 688; standard deviation = 153). Some Table VIII data reported in other previous research probably were

affected to some extent by truncation of range. For example, Hoffman (1989) reported that 95.5% of the crews at Grafenwoehr passed Table VIII (and hence their scores were confined to the interval between 700 and 1,000).

Future research should probe the influence of truncation of range on the M-COFT Hit Rate and Table VIII relationship. A first step would be to sample both variables from a population that could be expected to produce a relatively high Table VIII mean score. The correlation between the variables in this sample could be compared and contrasted with those in the present investigation, and then the data pooled to produce an omnibus test with maximum range on both variables.

Interestingly, truncation of range is the second methodological artifact that has been demonstrated to impact the relationship between Table VIII and M-COFT proficiency. The first artifact was suggested by Campshure and Drucker (1990) when they hypothesized that a composite measure of M-COFT achievement (e.g., M-COFT matrix position based on aggregated sessions, or a broad sample of M-COFT proficiency) may be necessary in order to reliably predict Table VIII scores. Table VIII performance represents a multi-faceted composite of many behaviors (including cognitive, motivational, and perceptual-motor functioning) as well as quality, extent and intensity of prior training. Because of the complexity of the criterion measure, Campshure and Drucker reasoned that only a composite sampling of M-COFT performance, encompassing a broad array of specific M-COFT behaviors, could reasonably be expected to predict Table VIII outcomes. Their hypothesis is supported by established tenets from psychometric theory, yet prior to their (1990) investigation, a great deal of research into the M-COFT-to-live-fire relationship was characterized by limited data samples on both the predictor and criterion sides of the prediction equation (e.g., Butler, et al., 1982; Kuma & McConville, 1982; Black & Abel, 1987; Hughes, et al., 1987).

Smith and Hagman (1992) confirmed the mitigating impact of insufficient samples of M-COFT proficiency. They reported that the relationship between M-COFT and Table VIII could be demonstrated convincingly only when M-COFT proficiency was indexed by the Hit Rate metric and Hit Rate was based upon a broad sample of M-COFT engagement conditions. Had they based their investigation on a single M-COFT exercise, they probably would have concluded that device proficiency bore little if any relationship to Table VIII scores. The present investigation confirmed the importance of a broad-based sample of M-COFT performance, and underscored the value of the Hit Rate metric.

Future research should explore which combinations of M-COFT exercises most reliably predict subsequent Table VIII performance. Although Hit Rate scores based on four exercises

predicted Table VIII scores in both the present research and in Smith and Hagman (1992), the vast majority of predictive utility in both investigations came from only two of the exercises. It is possible that a better combination of four exercises might be selected for the CTGP. In any event, care should be taken to avoid the pitfall of over-reliance on any single exercise. Because of the complexity of Table VIII performance, M-COFT prediction measures must encompass a broad array of specific simulation behaviors. Exercise Number 2 in the present investigation produced the highest correlation with Table VIII, but the same exercise in Smith and Hagman (1992) produced a negligible coefficient of correlation. For these reasons, it seems prudent to base future investigations of the M-COFT-to-live-fire relationship on the broadest possible sample of M-COFT exercises. Future research might probe the effect of expanding the sample base from four exercises to six, or even eight.

This admonition, that predictions of criterion measures should be based on the broadest possible sample of M-COFT proficiency, applies doubly when it comes to predicting Hit Rate based on its surrogate measure of proportion of first round kills. Although proportion of first round kills can be calculated on as little as one exercise (which requires about 15 min of training time in the M-COFT), results from both the present investigation and from Smith and Hagman (1992) suggest that one exercise is an insufficient sample of M-COFT proficiency. It is recommended that Hit Rate estimates should be based on M-COFT testing sessions that incorporate 40 or more targets, and that the engagements containing these targets should correspond in all important respects to conditions represented in Table VIII.

If these conditions are met, there can be little doubt that a reliable prediction model between M-COFT simulator scores and Table VIII scores will smooth the transition to device-based tank gunnery training strategies. RC commanders, by examining device proficiency, can gain an idea of whether crews are likely to qualify on Table VIII. If proportion of first round kills is based on an adequate sample of M-COFT targets, it is possible to use this metric as a substitute predictor of Hit Rate, with known bands of confidence around predicted scores. Thus, based only on the easily obtainable metric of proportion of first round kills, RC unit commanders can gain an idea of whether crews under their command will qualify on Table VIII. At the very least, Table 5 and its underlying model will permit identification of crews most in need of intensive remedial practice prior to Table VIII. Crews that are scoring 40% first round kills may qualify on Table VIII, but the odds are certainly against them. Crews shooting 50% of targets with first rounds have better odds of qualifying, and crews capable of hitting 61% of targets with first rounds within allocated time limits have a 50% chance of Table VIII qualification. Once crews rise above 61% of first round kills, their odds of Table VIII qualification increase steadily until,

at 93% of first round kills, their odds of Table VIII qualification stand at approximately 95%.

In conclusion, the findings of this research confirm that an M-COFT-to-Table VIII relationship exists when M-COFT proficiency is assessed by means of a composite measure like Hit Rate from the CTGP (Hoffman & Witmer, 1989) and Table VIII scores are widely dispersed. Based on this relationship, a tool is provided with which RC unit commanders can predict the probability of first-run crew qualification on Table VIII based on either Hit Rate or the easier to calculate percentage of first round kills obtained on the CTGP. Because of reasons mentioned earlier, data from future field implementation tryouts are still needed to verify the accuracy of these predictions. In the meantime, however, the findings of the present research represent a step forward in the development of a device-based tool for predicting RC live-fire tank gunnery success.

References

Black, B. A., & Abel, M. H. (1987). Review of U.S. armor crew and platoon training in preparation for the 1985 Canadian Army Trophy (CAT) competition (ARI Research Report 1442). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A185 470)

Black, B. A., & Graham, S. E. (1987). Armor crewmen assignment issues and the use of the Unit Conduct of Fire Trainer (U-COFT) in performance prediction (ARI Research Report 1458). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A192 062)

Butler, W. G., Reynolds, M. J., Kroh, M. Z., & Thorne, H. W. (1982). Training developments study--M1 (Abrams) Tank Unit-Conduct of Fire Trainer (TRASANA TEA 11-82). White Sands Missile Range, NM: U.S. Army TRADOC Systems Analysis Activity. (AD A173 938)

Campshure, D. A. (1991). Devices and aids for training M1 tank gunnery in the Army National Guard: A detailed analysis of training requirements (ARI Research Report 1588). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A240 931)

Campshure, D. A., & Drucker, E. H. (1990). Predicting first-run gunnery performance on tank table VIII (ARI Research Report 1571). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A228 201)

Du Bois, R. S. (1987). The M1 Unit-Conduct of Fire Trainer (U-COFT) as a tank gunnery testing device: A psychometric evaluation. Unpublished master's thesis, Western Kentucky University, Bowling Green, KY.

Graham, S. E. (1986). The Unit-Conduct of Fire Trainer (U-COFT) as a medium for assessing gunner proficiency: Test reliability and utility (ARI Research Report 1422). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A196 196)

Hays, W. L. (1963). Statistics. New York: Holt, Rinehart and Winston.

Hoffman, R. G. (1989). Description and prediction of Grafenwoehr M1 tank table VIII performance (ARI Technical Report 837). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B136 331)

Hoffman, R. G., & Witmer, B. G. (1989). Development of a Unit-Conduct of Fire Trainer (U-COFT) test of M1 gunnery proficiency (ARI Technical Report 859). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A219 045)

Hughes, C. R., Butler, W. G., Sterling, B. S., & Bergland, A. W. (1987). M1 Unit Conduct-of-Fire Trainer post fielding training effectiveness analyses (TRAC-WSMR-TEA-16-87). White Sands Missile Range, NM: U.S. Army TRADOC Analysis Command. (AD B113 298)

Kuma, D., & McConville, L. (1982). Independent evaluation report for M1/M60 series Unit Conduct of Fire Trainer (U-COFT) (TRADOC ACN 39373). Fort Knox, KY: U.S. Army Armor Center, Directorate of Training Developments. (AD B065 441)

Morrison, J. E., Campshure, D. A., & Doyle, E. L. (1991). A Device/Aid-Based Strategy for training M1 tank gunnery in the Army National Guard (ARI Research Report 1587). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A240 752)

Morrison, J. E., Drucker, E. H., & Campshure, D. A. (1991). Devices and aids for training M1 tank gunnery in the Army National Guard: A review of military documents and the research literature (ARI Research Report 1586). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A240 628)

Smith, M. D., & Hagman, J. D. (1992). Predicting Table VIII tank gunnery performance from M-COFT hit rate and demographic variables (ARI Technical Report 955). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A254 580)

U.S. Army Armor Center (1985). M1 Unit Conduct-of-Fire Trainer (U-COFT) training device support package (FC 17-12-7-1). Fort Knox, KY.

U.S. Department of the Army. (1988, September). Tank combat tables M1 (C2, FM 17-12-1). Washington, DC. (Available from the Commander, U.S. Army Armor Center and Fort Knox, ATTN: ATZK-PTF-D, Fort Knox, KY 40121-5200.)